The authors thank reviewer 2 for his/her second round of constructive and relevant comments. The authors also thank the editor and reviewer 1 (Clement Albergel) for their diligent handling of the paper. The comments are shown in ***italic bold fonts***, our responses are in regular fonts, and adjusted text from the manuscript is marked with "Correction:".

Reviewer 2

Main Comments:

1) I am still a bit lost about the purpose of the paper, likely because there are multiple interacting components of the goals. It seems there are (1) very broad goals related to generally improving NASA LIS parametrizations and (2) there are very specific and narrow goals about studying Dry Chaco water balance and effects of deforestation. Both are fine by themselves, but these are currently mixed together in this paper (for example in Lines 85-90) and not easy to follow throughout the paper. This mixing seems to confound each analysis. The general goal (1) is limited by the fact that the Dry Chaco appears to be a not well-behaving location (has large land use change). Therefore, it could be a strange location to make general LSM improvements and assess LSM differences as opposed to a different location that does not have land use change confounding factors. The specific goals (2) initially appear confounded by the general goals because the authors do not mention that the deforestation effect and vegetation parameters would be fully isolated (which they are in the methods) from investigations with SHPs.

There are a lot of changes being made: vegetation parameters, SHPs, changing LSMs, etc. The authors either need to focus on one (an extreme change would be to separate the broad and specific goals into different papers). Or they need to clearly partition each goal in the introduction and what the current problem is in the field with each. The paper ends with a list of conclusions on Line 655 that don't feel very organized in addressing the multifaceted goals. I recommend acknowledging these differences in the broad and narrow goals and be very clear why the Dry Chaco is the best testbed to evaluate all of these interacting goals.

These ideas should be clearer in lines 85-90 and throughout, including the conclusions. Lines 1-5 in the abstract are close in doing that: the authors start with the big picture to make the overall LIS improvement and narrow into the specific deforestation issues. This is mostly good. But I am still left wondering whether the Dry Chaco is the best location to study the overall water balance and test overall LSM performance if it is undergoing major land use change that could confound tests with those parameters.

Many thanks for the new recommendations and clarifications on the confusion. We agree that there are multiple interacting objectives in our study that could be further clarified. The idea of splitting the objectives into general and narrow goals is highly appreciated and implemented in the updated manuscript. The conclusion-section was revised following the same structure. Also the reasoning behind the different hypotheses was further elaborated. The corrections are summarized below by responding to the specific comments.

I still find limited context in the introduction about the state of the knowledge in the model performance. For example, the statement starting on line 29 leaves the reader wondering: how well do LSMs currently represent land surface processes and where are the knowledge

gaps and model limitations? It is a nice statement to set the rest of the introduction up. But afterwards, the wording is vague and sounds a bit obvious: better parameterizations will make the model better. There is great discussion in the Discussion section on the state of the field, so it seems the motivation is there. The authors should be more explicit up front about this. I think a new paragraph needs to address, for example: have previous studies showed specific issues with SHP parameters that caused large biases/RMSE? What specifically happened in previous studies when proper vegetation change and phenology was not included? These are just examples, but more direct motivating statements answering such questions would strengthen the motivation of the paper.

Thanks for the suggestion. We agree that some parts of the introduction were a bit vague and that the motivations behind the research could be improved. Following sentences were added or improved, and we explicitly refer to specific studies that address how SHP or vegetation parameters are connected to model performance.

Correction
Line 49: Earlier studies indicated that replacing climatological vegetation by interannually varying satellite-derived indices can improve modeled energy fluxes as well as surface temperature and moisture in both offline LSM simulations (Miller et al., 2006; Case et al., 2013; Yin et al., 2016) and atmosphere-coupled LSMs (Crawford et al., 2001; James et al., 2009; Boussetta et al., 2013; Ge et al., 2014; Kumar et al., 2014). The largest improvements are obtained during extreme meteorological anomalies (Case et al., 2013). In this study, it is expected that besides meteorological anomalies, also large-scale land cover conversions, such as deforestation, alter the vegetation strongly from its climatological representation. Therefore, it is tested if the use of satellite-derived vegetation indices in LSMs is also gainful in regions characterized by land cover changes.

Line 60: The sensitivity of LSMs to plant functional types or land cover related parameters has been illustrated in both offline (Chen et al., 2014) and atmosphere-coupled LSMs (Pitman et al., 2009; Grossman-Clarke et al., 2010; Cao et al., 2015; Ruiz-Vasquez et al., 2020). Most of these studies solely focused on changes in model-specific surface parameters without taking into account changes in LAI or GVF. In contrast, this study aims at implementing large-scale land cover changes in LSMs by feeding them with both temporally varying vegetation indices and land cover parameters.

Line 66: Different LSMs have different soil parameterizations and use model-specific, often historically tuned, SHPs.

Line 71: As shown by De Lannoy et al. (2014), the implementation of more accurate soil texture and related SHPs, can lead to reduced bias and error-estimates of soil moisture when compared to in situ surface soil moisture and even impact simulated runoff and evaporation estimates.

Specific Comments:

Line 9: large regional and long term differences in what?

Correction:
Line 9: A relative comparison in terms of water budget components and ‘efficiency space’ for various baseline and revised experiments showed that large regional and long-term

differences in the simulated water budget partitioning relate to different LSM structures, whereas smaller local differences resulted from updated soil, vegetation and land cover parameters.

Line 72-75: Since I am not very clear on the goals, I am a bit lost with this hypothesis. First of all, the hypothesis here seems obvious – that it would have to be true. Aren't the authors, in part, trying to evaluate only the vegetation to evaluate its effect in representing the deforestation? Then separately evaluate the soil component to see how this additionally improves it? So the hypothesis would be something about the fact that improving vegetation parameterizations greatly improves model improvement in the Dry Chaco over effect of SHP updates.

Correction related to hypotheses:
Line 85: Given the Dry Chaco's recent and large-scale deforestation history, it is a unique area to test the impact of temporally evolving vegetation and land cover parameters. It is expected that by feeding LSMs with time-varying vegetation and land cover, together with an updated set of soil parameters, the most accurate spatial and temporal representation of the Chaco's water distribution could be obtained. Additionally, it is hypothesized that the use of similar soil, vegetation and land cover parameters in various LSMs, would result in similar accurate estimates of the long-term simulated water budgets. Lastly, it is hypothesized that soil and vegetation parameter updates would contribute differently towards the model performance improvement.

Correction related to the objectives:
Line 97: Besides these general objectives, the simulated water budget components are further framed within the hydrological context of the Dry Chaco, i.e. it is verified if the different LSMs simulate the increased deep percolation and higher soil moisture values after deforestation, similar to the field-based findings of Nosetto et al. (2012), Giménez et al. (2016), Magliano et al. (2017) and Marchesini et al. (2017).

Line 78: Wouldn't one want to evaluate only vegetation parameter enhancements to isolate the enhancement with deforestation? On the other hand, wouldn't one want to evaluate improvement in soil parameters in a region where no deforestation is occurring? I would make it clear here that the methods here isolate the effects of each. It almost sounds like the study runs the risk of making multiple changes at once and not being able to attribute which changes are truly enhancing the model – which is not the case here.

Good point, we added an extra paragraph on the general set-up of the paper, clarifying the structure of the methodology and result-section.

Correction:
Line 101: The impact of the revised set of soil parameters and updated vegetation and land cover treatment is analyzed incrementally. In a first phase, models were run with their default model-specific soil parameters, climatological vegetation (LAI and GVF) and static land cover. Next, the models were supplied with more accurate soil texture and related SHPs and their impact on the simulated water budgets was quantified. In a third phase, the ongoing land cover changes were implemented using interannually varying satellite-based indices and yearly updated land cover maps and it was analyzed how the major land cover changes alter

the hydrological balance. Lastly, the impact of the various model structures, soil texture and dynamic vegetation input was assessed using the concept of 'efficiency space' and the performance of each set of experiments was evaluated against independent satellite-based estimates of evapotranspiration, brightness temperature and in situ soil moisture.

Line 85: It still isn't clear to me: is the Dry Chaco the best region to evaluate all changes being made (different LSMs, SHPs, and vegetation parameters)? It seems clear to evaluate the effect of deforestation here. But there are a bunch of other goals that make me thing the Dry Chaco is not the best for. Wouldn't one want to test the effect of different SHPs and LSMs at a site that does not have a strong land use change signal? Please clearly list all the reasons why the Dry Chaco was chosen for this multifaceted study (probably in Section 2.1) – the deforestation is good for some parts, but seems counterproductive for other parts of the study.

We hope that the extra clarifications on the hypotheses and objectives of our research make clear now why the Dry Chaco was selected.

Line 91: It would be helpful to specifically mention what each of these datasets is being used for each of subsections of Section 2. For example, is the data used for model input or validation?

We added a new subsection '2.3 Revised input data', in which the updated input parameter sets are presented. This similar to section 2.4 (Evaluation data), where the evaluation datasets are presented.

Lines 313-318: Efficiency curves are a holistic and meaningful way to compare between models. However, the language in these lines makes it unclear what the goals are and what the reader is looking for in Figs 8 and 9. It seems the authors desire in-situ data to see if the updated SHPs represent in-situ behavior better, but it is not available. Line 316 then says that a comparison is being made, but it needs to be clearer what the authors are testing. Is it to find which LSM performance seems "better"? Is it to evaluate the hydrologic behavior itself that the models provide some consensus on and how updated SHPs change this behavior?

Correction:
Line 346: The efficiency plots are used to see if the three LSMs provide a consensus on the simulated hydrological behavior and on the impact of the updated soil, vegetation and land cover treatment.

Line 481: This 4.3.1 section can seem distracting at first. At this point of the paper, I am not sure why one needs to know whether MERRA2 precipitation compares to in-situ measurements in the context of the study goals. One finds out the purpose in line 603. It should be explained earlier why this analysis is being done and how a P bias can confound the results.

Correction:
Line 502: Because the quality of input P will greatly influence the quality of ET, sfmc and Tb simulations, independent of the used LSM, the quality of the MERRA2 P product was first evaluated against in situ P data.

Line 646: Is the study about the Dry Chaco or the LSMs? This statement seems to be focused more on the Dry Chaco hydrology, but the study is mainly focused on LSMs and not Dry Chaco itself (the main findings bullets in the conclusion here all focus on the LSM parametrizations and not Dry Chaco). As I mentioned earlier, the conclusions appear to be a bit unorganized with respect to which of the objectives each is addressing. If the goals are more clearly structured and outlined, it will be easier to follow how each conclusion addressed each goal.

Thanks for the remarks. The conclusion was restructured:

Correction:
Line 669: In this study, we updated the soil- and vegetation-related parameters of three LSMs (CLM2.0, CLSM-F2.5 and Noah3.6), grouped within NASA's LIS, to obtain the best modeled representation of the hydrology over the South-American Dry Chaco. We used HWSD v1.21 soil texture and time-varying satellite-based GLASS and GIMMS vegetation indices, along with yearly updated ESA-CCI land cover information. The impact of the various model structures, soil texture and dynamic vegetation input was assessed in terms of water budget partitioning and efficiency space. Our results indicate that:

- the three LSMs yield a different partitioning of the water budget, with 74% to 95% of the total annual P over the Dry Chaco contributing to ET;
- the soil texture pattern is the main driver of the spatial pattern of soil moisture;
- introducing similar soil, vegetation and land cover parameters in the various LSMs does not result in a homogenization of the long-term water budget components, i.e. the various LSM structures primarily determine the water distribution whereas soil, vegetation and land cover parameters only have a secondary impact.

The updated vegetation and land cover treatment allowed to explore to which extent large-scale land cover changes in the Dry Chaco affect the different water budget components. It was found that:

- deforestation increases soil moisture for all LSMs, but the degree of increase is depending on the model structure;
- a change in land cover results in a shift of the model climatology and a (non-stationary) redistribution of the water budget, which is different for each LSM;
- the implemented satellite-based vegetation indices do not fully depict deforestation, because the 0.125° spatial resolution partially suppresses the deforestation signal, and the replacing agricultural crop may have similar LAI and GVF values as the initial dry forest.

The model input and output were further evaluated against independent data of in situ P and sfmc, and spatially covering GLEAM ET and SMOS Tb. The latter offers the unique possibility for an integral evaluation of simulated soil moisture, soil temperature and LAI, after forwarding these variables through a zero-order RTM. Relative to independent data, no specific LSM structure, soil or vegetation input is significantly better than another in terms of time series metrics …